

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/424,951

DATE: 03/05/2001

TIME: 15:43:55

Input Set : A:\87714113.app
Output Set: N:\CRF3\03052001\I424951.raw

3 <110> APPLICANT: SRIKANTHA, THYAGARAJAN
4 SOLL, DAVID R.
6 <120> TITLE OF INVENTION: CANDIDA ALBICANS TWO-COMPONENT HYBRID KINASE GENE,
7 CANIK1, AND USE THEREOF
9 <130> FILE REFERENCE: 087714/0113
11 <140> CURRENT APPLICATION NUMBER: 09/424,951
12 <141> CURRENT FILING DATE: 2000-01-20
14 <150> PRIOR APPLICATION NUMBER: PCT/US98/11658
15 <151> PRIOR FILING DATE: 1998-06-05
17 <150> PRIOR APPLICATION NUMBER: 60/048,914
18 <151> PRIOR FILING DATE: 1997-06-06
20 <160> NUMBER OF SEQ ID NOS: 7
22 <170> SOFTWARE: PatentIn Ver. 2.1
24 <210> SEQ ID NO: 1
25 <211> LENGTH: 1254
26 <212> TYPE: DNA
27 <213> ORGANISM: Candida albicans
29 <220> FEATURE:
30 <221> NAME/KEY: CDS
31 <222> LOCATION: (1)..(1254)
33 <400> SEQUENCE: 1
34 gag att aga aca cca ttg aat ggg att att ggw atg acy cag ttg tcr 48
35 Glu Ile Arg Thr Pro Leu Asn Gly Ile Ile Gly Met Thr Gln Leu Ser
36 1 5 10 15
38 ctt gat aca gag ttg acr cag tac caa cga gag atg ttg tcg att gtg 96
39 Leu Asp Thr Glu Leu Thr Gln Tyr Gln Arg Glu Met Leu Ser Ile Val
40 20 25 30
42 cat aac ttg gca aat tcc ttg ttg acc att ata gac gat ata ttg gat 144
43 His Asn Leu Ala Asn Ser Leu Leu Thr Ile Ile Asp Asp Ile Leu Asp
44 35 40 45
46 att tct aag att gag gcg aat aga atg acg gtg gaa cag att gat ttt 192
47 Ile Ser Lys Ile Glu Ala Asn Arg Met Thr Val Glu Gln Ile Asp Phe
48 50 55 60
50 tca tta aga ggg aca gtg ttt ggt gca ttg aaa acg tta gcc gtc aaa 240
51 Ser Leu Arg Gly Thr Val Phe Gly Ala Leu Lys Thr Leu Ala Val Lys
52 65 70 75 80
54 gct att gaa aaa aac cta gac ttg acc tat caa tgt gat tca tcg ttt 288
55 Ala Ile Glu Lys Asn Leu Asp Leu Thr Tyr Gln Cys Asp Ser Ser Phe
56 85 90 95
58 cca gat aat ctt att gga gat agt ttt aga tta cga caa gtt att ctt 336
59 Pro Asp Asn Leu Ile Gly Asp Ser Phe Arg Leu Arg Gln Val Ile Leu
60 100 105 110
62 aac ttg gct ggt aat gct att aag ttt act aaa gag ggg aaa gtt agt 384
63 Asn Leu Ala Gly Asn Ala Ile Lys Phe Thr Lys Glu Gly Lys Val Ser
64 115 120 125
66 gtt agt gtg aaa aag tct gat aaa atg gtg tta gat agt aag ttg ttg 432
67 Val Ser Val Lys Lys Ser Asp Lys Met Val Leu Asp Ser Lys Leu Leu

ENTERED

See p5

BEST AVAILABLE COPY

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/424,951

DATE: 03/05/2001
TIME: 15:43:55

Input Set : A:\87714113.app
Output Set: N:\CRF3\03052001\I424951.raw

68	130	135	140	
70	tta gag gtt tgt gtt agc gac acg gga ata ggt ata gag aaa gac aaa			480
71	Leu Glu Val Cys Val Ser Asp Thr Gly Ile Gly Ile Glu Lys Asp Lys			
72	145	150	155	160
74	ttg gga ttg att ttc gat acc ttc tgt caa gct gat ggt tct act aca			528
75	Leu Gly Leu Ile Phe Asp Thr Phe Cys Gln Ala Asp Gly Ser Thr Thr			
76	165	170	175	
78	aga aag ttt ggt ggt aca ggt tta ggg ttg tca att tcc aaa cag ttg			576
79	Arg Lys Phe Gly Gly Thr Gly Leu Gly Leu Ser Ile Ser Lys Gln Leu			
80	180	185	190	
82	ata cat tta atg ggt gga gag ata tgg gtt act tcg gag tat gga tcc			624
83	Ile His Leu Met Gly Gly Glu Ile Trp Val Thr Ser Glu Tyr Gly Ser			
84	195	200	205	
86	ggr tca aac ttt tat ttt acg gtg tgc gtg cca tct aat att aga			672
87	Gly Ser Asn Phe Tyr Phe Thr Val Cys Val Ser Pro Ser Asn Ile Arg			
88	210	215	220	
90	tat act cga caa acc gaa caa ttg tta cca ttt agt tcc cat tat gtg			720
91	Tyr Thr Arg Gln Thr Glu Gln Leu Leu Pro Phe Ser Ser His Tyr Val			
92	225	230	235	240
94	tta ttt gta tcg act gag cat act caa gaa ctt gat gtg ttg aga			768
95	Leu Phe Val Ser Thr Glu His Thr Gln Glu Leu Asp Val Leu Arg			
96	245	250	255	
98	gat gga att ata gaa ctt gga ttg ata cct ata ata gtg aga aat att			816
99	Asp Gly Ile Ile Glu Leu Gly Leu Ile Pro Ile Ile Val Arg Asn Ile			
100	260	265	270	
102	gaa gat gca aca ttg act gag ccg gtg aaa tat gat ata att atg att			864
103	Glu Asp Ala Thr Leu Thr Glu Pro Val Lys Tyr Asp Ile Ile Met Ile			
104	275	280	285	
106	gat tcg ata gag att gcc aaa aag ttg agg ttg tta tcg gag gtt aaa			912
107	Asp Ser Ile Glu Ile Ala Lys Lys Leu Arg Leu Leu Ser Glu Val Lys			
108	290	295	300	
110	tat att ccg ttg gtt ttg gtc cat cat tct att cca cag ttg aat atg			960
111	Tyr Ile Pro Leu Val Leu Val His His Ser Ile Pro Gln Leu Asn Met			
112	305	310	315	320
114	aga gta tgt att gat ttg ggg ata tct tcc tat gca aat acg cca tgt			1008
115	Arg Val Cys Ile Asp Leu Gly Ile Ser Ser Tyr Ala Asn Thr Pro Cys			
116	325	330	335	
118	tcg atc acg gac ttg gcc agt gcg att ata cca gcg ttg gag tcg aga			1056
119	Ser Ile Thr Asp Leu Ala Ser Ala Ile Ile Pro Ala Leu Glu Ser Arg			
120	340	345	350	
122	tct ata tca cag aac tca gac gag tcg gtg agg tac aaa ata tta cta			1104
123	Ser Ile Ser Gln Asn Ser Asp Glu Ser Val Arg Tyr Lys Ile Leu Leu			
124	355	360	365	
126	gca gag gac aac ctc gtc aat cag aaa ctt gca gtt agg ata tta gaa			1152
127	Ala Glu Asp Asn Leu Val Asn Gln Lys Leu Ala Val Arg Ile Leu Glu			
128	370	375	380	
130	aag caa ggg cat ctg gtg gaa gta gtt gag aac gga ctc gag gcg tac			1200
131	Lys Gln Gly His Leu Val Glu Val Val Glu Asn Gly Leu Glu Ala Tyr			
132	385	390	395	400

TECH CENTER 1600/2900

MAR 12 2001

RECEIVED

BEST AVAILABLE COPY

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/424,951

DATE: 03/05/2001
TIME: 15:43:55

Input Set : A:\87714113.app
Output Set: N:\CRF3\03052001\I424951.raw

134 gaa gcg att aag agg aat aaa tat gat gtg gtg ttg atg gat gtg caa 1248
 135 Glu Ala Ile Lys Arg Asn Lys Tyr Asp Val Val Leu Met Asp Val Gln
 136 405 410 415
 138 atg cct 1254
 139 Met Pro
 142 <210> SEQ ID NO: 2
 143 <211> LENGTH: 418
 144 <212> TYPE: PRT
 145 <213> ORGANISM: Candida albicans
 147 <400> SEQUENCE: 2
 148 Glu Ile Arg Thr Pro Leu Asn Gly Ile Ile Gly Met Thr Gln Leu Ser
 149 1 5 10 15
 151 Leu Asp Thr Glu Leu Thr Gln Tyr Gln Arg Glu Met Leu Ser Ile Val
 152 20 25 30
 154 His Asn Leu Ala Asn Ser Leu Leu Thr Ile Ile Asp Asp Ile Leu Asp
 155 35 40 45
 157 Ile Ser Lys Ile Glu Ala Asn Arg Met Thr Val Glu Gln Ile Asp Phe
 158 50 55 60
 160 Ser Leu Arg Gly Thr Val Phe Gly Ala Leu Lys Thr Leu Ala Val Lys
 161 65 70 75 80
 163 Ala Ile Glu Lys Asn Leu Asp Leu Thr Tyr Gln Cys Asp Ser Ser Phe
 164 85 90 95
 166 Pro Asp Asn Leu Ile Gly Asp Ser Phe Arg Leu Arg Gln Val Ile Leu
 167 100 105 110
 169 Asn Leu Ala Gly Asn Ala Ile Lys Phe Thr Lys Glu Gly Lys Val Ser
 170 115 120 125
 172 Val Ser Val Lys Lys Ser Asp Lys Met Val Leu Asp Ser Lys Leu Leu
 173 130 135 140
 175 Leu Glu Val Cys Val Ser Asp Thr Gly Ile Gly Ile Glu Lys Asp Lys
 176 145 150 155 160
 178 Leu Gly Leu Ile Phe Asp Thr Phe Cys Gln Ala Asp Gly Ser Thr Thr
 179 165 170 175
 181 Arg Lys Phe Gly Gly Thr Gly Leu Gly Leu Ser Ile Ser Lys Gln Leu
 182 180 185 190
 184 Ile His Leu Met Gly Gly Glu Ile Trp Val Thr Ser Glu Tyr Gly Ser
 185 195 200 205
 187 Gly Ser Asn Phe Tyr Phe Thr Val Cys Val Ser Pro Ser Asn Ile Arg
 188 210 215 220
 190 Tyr Thr Arg Gln Thr Glu Gln Leu Leu Pro Phe Ser Ser His Tyr Val
 191 225 230 235 240
 193 Leu Phe Val Ser Thr Glu His Thr Gln Glu Glu Leu Asp Val Leu Arg
 194 245 250 255
 196 Asp Gly Ile Ile Glu Leu Gly Leu Pro Ile Ile Val Arg Asn Ile
 197 260 265 270
 199 Glu Asp Ala Thr Leu Thr Glu Pro Val Lys Tyr Asp Ile Ile Met Ile
 200 275 280 285
 202 Asp Ser Ile Glu Ile Ala Lys Lys Leu Arg Leu Leu Ser Glu Val Lys
 203 290 295 300
 205 Tyr Ile Pro Leu Val Leu Val His His Ser Ile Pro Gln Leu Asn Met

BEST AVAILABLE COPY

RAW SEQUENCE LISTING DATE: 03/05/2001
 PATENT APPLICATION: US/09/424,951 TIME: 15:43:55

Input Set : A:\87714113.app
 Output Set: N:\CRF3\03052001\I424951.raw

206	305	310	315	320
208	Arg Val Cys Ile Asp Leu Gly Ile Ser Ser Tyr Ala Asn Thr Pro Cys			
209	325	330	335	
211	Ser Ile Thr Asp Leu Ala Ser Ala Ile Ile Pro Ala Leu Glu Ser Arg			
212	340	345	350	
214	Ser Ile Ser Gln Asn Ser Asp Glu Ser Val Arg Tyr Lys Ile Leu Leu			
215	355	360	365	
217	Ala Glu Asp Asn Leu Val Asn Gln Lys Leu Ala Val Arg Ile Leu Glu			
218	370	375	380	
220	Lys Gln Gly His Leu Val Glu Val Val Glu Asn Gly Leu Glu Ala Tyr			
221	385	390	395	400
223	Glu Ala Ile Lys Arg Asn Lys Tyr Asp Val Val Leu Met Asp Val Gln			
224	405	410	415	
226	Met Pro			
230	<210> SEQ ID NO: 3			
231	<211> LENGTH: 3246			
232	<212> TYPE: DNA			
233	<213> ORGANISM: Candida albicans			
235	<220> FEATURE:			
236	<221> NAME/KEY: CDS			
237	<222> LOCATION: (1)..(3243)			
239	<400> SEQUENCE: 3			
240	atg aac ccc act aaa aaa cct cgg tta tca cca atg cag ccc tct gtt	48		
241	Met Asn Pro Thr Lys Pro Arg Leu Ser Pro Met Gln Pro Ser Val			
242	1 5 10 15			
244	ttt gaa ata ctc aac gac cct gag ctt tat agt cag cac tgt cat agc	96		
245	Phe Glu Ile Leu Asn Asp Pro Glu Leu Tyr Ser Gln His Cys His Ser			
246	20 25 30			
248	ctt agg gaa aca ctt ctt gat cat ttc aac cat caa gct aca ctt atc	144		
249	Leu Arg Glu Thr Leu Leu Asp His Phe Asn His Gln Ala Thr Leu Ile			
250	35 40 45			
252	gac act tat gaa cat gaa cta gaa aaa tcc aaa aat gcc aac aaa gcg	192		
253	Asp Thr Tyr Glu His Glu Leu Glu Lys Ser Lys Asn Ala Asn Lys Ala			
254	50 55 60			
256	tcc caa caa gca ctt agt gaa ata ggt aca gtt gtt ata tct gtt gcc	240		
257	Ser Gln Gln Ala Leu Ser Glu Ile Gly Thr Val Val Ile Ser Val Ala			
258	65 70 75 80			
260	atg gga gac ttg tcg aaa aaa gtt gag att cac aca gta gaa aat gac	288		
261	Met Gly Asp Leu Ser Lys Val Glu Ile His Thr Val Glu Asn Asp			
262	85 90 95			
264	cct gag att tta aaa gtc aaa atc acc atc aac acc atg atg gat caa	336		
265	Pro Glu Ile Leu Lys Val Lys Ile Thr Ile Asn Thr Met Met Asp Gln			
266	100 105 110			
268	tta cag aca ttt gct aat gag gtt aca aaa gtc gcc acc gaa gtc gca	384		
269	Leu Gln Thr Phe Ala Asn Glu Val Thr Lys Val Ala Thr Glu Val Ala			
270	115 120 125			
272	aat ggt gaa cta ggt gga caa gcg aaa aat gat gga tct gtt ggt att	432		
273	Asn Gly Glu Leu Gly Gly Gln Ala Lys Asn Asp Gly Ser Val Gly Ile			
274	130 135 140			

BEST AVAILABLE COPY

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/424,951

DATE: 03/05/2001
TIME: 15:43:56

Input Set : A:\87714113.app
Output Set: N:\CRF3\03052001\I424951.raw

276	tgg	aga	tca	ctt	aca	gac	aat	gtt	aat	att	atg	gct	ctt	aat	tta	act	480	
277	Trp	Arg	Ser	Leu	Thr	Asp	Asn	Val	Asn	Ile	Met	Ala	Leu	Asn	Leu	Thr		
278	145									150						155	160	
280	aac	caa	gtg	cga	gaa	att	gct	gat	gtc	aca	cgt	gct	gtt	gcc	aag	ggg	528	
281	Asn	Gln	Val	Arg	Glu	Ile	Ala	Asp	Val	Thr	Arg	Ala	Val	Ala	Lys	Gly		
282										165						170	175	
284	gac	ttg	tca	cgt	aaa	att	aat	gta	cac	gcc	cag	ggt	gaa	atc	ctt	caa	576	
285	Asp	Leu	Ser	Arg	Lys	Ile	Asn	Val	His	Ala	Gln	Gly	Glu	Ile	Leu	Gln		
286										180						185	190	
288	ctt	caa	cgt	aca	ata	aac	acc	atg	gtg	gat	cag	tta	cga	acg	ttt	gca	624	
289	Leu	Gln	Arg	Thr	Ile	Asn	Thr	Met	Val	Asp	Gln	Leu	Arg	Thr	Phe	Ala		
290										195						200	205	
292	tcc	gaa	gta	tct	aaa	gtt	gct	aga	gat	gtt	gtg	ctt	ggt	ata	tta	672		
293	Phe	Glu	Val	Ser	Lys	Val	Ala	Arg	Asp	Val	Gly	Val	Leu	Gly	Ile	Leu		
294										210						215	220	
296	gga	gga	caa	gcg	ttg	att	gaa	aat	gtt	gaa	ggt	att	tgg	gaa	gag	ttg	720	
297	Gly	Gly	Gln	Ala	Leu	Ile	Glu	Asn	Val	Glu	Gly	Ile	Trp	Glu	Glu	Leu		
298										225						230	235	240
300	act	gat	aat	gtc	aat	gcc	atg	gct	ctt	aat	ttg	act	aca	caa	gtg	aga	768	
301	Thr	Asp	Asn	Val	Asn	Ala	Met	Ala	Leu	Asn	Leu	Thr	Thr	Gln	Val	Arg		
302										245						250	255	
304	aat	att	gcc	aat	gtc	acc	act	gcc	gtt	gcc	aag	ggg	gat	ttg	tcg	aaa	816	
305	Asn	Ile	Ala	Asn	Val	Thr	Thr	Ala	Val	Ala	Lys	Gly	Asp	Leu	Ser	Lys		
306										260						265	270	
308	aaa	gtc	act	gct	gat	tgt	aag	gga	gaa	aty	ctt	gat	ttg	aaa	ctt	act	864	
309	Lys	Val	Thr	Ala	Asp	Cys	Lys	Gly	Glu	Ile	Leu	Asp	Leu	Lys	Leu	Thr		
310										275						280	285	
312	att	aat	caa	atg	gtg	gac	cga	tta	cag	aat	ttt	gct	ctt	gcg	gtg	acg	912	
313	Ile	Asn	Gln	Met	Val	Asp	Arg	Leu	Gln	Asn	Phe	Ala	Leu	Ala	Val	Thr		
314										290						295	300	
316	aca	ttg	tcg	aga	gag	gtt	ggt	act	ttg	ggt	att	ttg	ggt	gga	caa	gct	960	
317	Thr	Leu	Ser	Arg	Glu	Val	Gly	Thr	Leu	Gly	Ile	Leu	Gly	Gly	Gln	Ala		
318										305						310	315	320
320	aac	gta	cag	gat	gtt	gaa	ggt	gct	tgg	aaa	cag	gtt	aca	gaa	aat	gtc	1008	
321	Asn	Val	Gln	Asp	Val	Glu	Gly	Ala	Trp	Lys	Gln	Val	Thr	Glu	Asn	Val		
322										325						330	335	
324	aac	cta	atg	gct	act	aat	tta	act	aac	caa	gtg	aga	tct	att	gct	aca	1056	
325	Asn	Leu	Met	Ala	Thr	Asn	Leu	Thr	Asn	Gln	Val	Arg	Ser	Ile	Ala	Thr		
326										340						345	350	
328	gtt	act	act	gca	gtt	gcg	cat	ggt	gat	ttg	tcg	caa	aag	att	gat	ggt	1104	
329	Val	Thr	Thr	Ala	Val	Ala	His	Gly	Asp	Leu	Ser	Gln	Lys	Ile	Asp	Gly		
330										355						360	365	
332	cat	ccc	aaa	gga	gag	att	tta	caa	ttg	aaa	aat	aca	atc	aac	aag	atg	1152	
333	His	Pro	Lys	Gly	Glu	Ile	Leu	Gln	Leu	Lys	Asn	Thr	Ile	Asn	Lys	Met		
334										370						375	380	
336	gtg	gac	tct	ttg	cag	ttg	ttt	gca	tca	gaa	gtg	tcg	aaa	gtg	gca	caa		
337	Val	Asp	Ser	Leu	Gln	Leu	Phe	Ala	Ser	Glu	Val	Ser	Lys	Val	Ala	Gln		
338										385						390	395	400
340	gat	gtt	ggt	att	aat	gga	aaa	tta	ggt	att	caa	gca	caa	gtt	agt	gat	1248	

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one 'n' or Xaa.

BEST AVAILABLE COPY

VERIFICATION SUMMARY
PATENT APPLICATION: US/09/424,951

DATE: 03/05/2001
TIME: 15:43:57

Input Set : A:\87714113.app
Output Set: N:\CRF3\03052001\I424951.raw

L:743 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5
L:760 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6
L:777 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7

BEST AVAILABLE COPY